Does Evolution cause a Domain Shift?

Konstantinos Rematas, Basura Fernando, Tatiana Tommasi, Tinne Tuytelaars
KU Leuven, ESAT-PSI, iMinds, Belgium
{first_name.last_name}@esat.kuleuven.be

1. Introduction

Domain adaptation methods aim at building classifiers that are robust to mismatched distributions in training and test. In computer vision such a mismatch can have several causes. Some are related to the image acquisition process like differences in the camera quality, variations in background, viewpoint and illumination conditions [7]. Others depend on higher level changes like moving from painting or sketches to real pictures of the same object or scene [8]. One further possible cause of domain shift that has not been studied yet is the design variation over time. For instance, think about what we call *telephone* today and its evolution in terms of visual aspect since the beginning of the 20th century (Figure 1, left). The same happens for fine-grained object categories like *cars*. Many contemporary car models are the successors of models that were introduced decades ago and still keep the same name, despite the change in appearance (Figure 1, right). Here we focus on this domain shift and we analyze how the state of the art domain adaptation methods perform over it. For this purpose we define a new testbed by collecting a *car evolution dataset* [1].

Figure 1. Design variation over time of the telephone (left) and the Mercedes E Class (right).

2. Car Evolution Dataset

Although the car class exists in many computer vision datasets, the information over the production year is not provided. The samples are usually organized according to the car type, namely sport car, sedan, minivan etc. (*e.g.* in [3]). In contrast, we keep the timeline as a separate axis and we identify any instance of the category car on the basis of its manufacturer, its model and its production year. Car models that belong to consecutive years have usually small differences but the extremes of the model's lifespan are considerably different. An analogous consideration moved the work of Y. J. Lee [6], which was developed simultaneously and independently from ours. In that work the authors focus on style change along the years and propose a method to estimate the production date of a car regardless of its model or manufacturer.

We defined our dataset by collecting images from Google Image Search API; the search term consisted of the manufacturer, the model and the line (*e.g.* Mercedes S Class W116). The line term corresponds to a particular production year and we manually created the correspondences from lines to years in the range 1972-2013. The downloaded images were then manually cleaned by removing all the samples depicting the interior of the cars and other noisy data. The process resulted in approximately 30 images per car line and 1088 images in total. The obtained hierarchy is shown in Figure 2.

3. Experiments

For our experiments we consider two classification tasks, the first based on the car manufacturer (3 classes) and the second on the car model (6 classes). We group the images on the basis of the car production decades 1970, 1980, 1990 and 2000 and we use the first three decades as the source and the last one as the target domain. We also divide the source samples into two roughly equal parts respectively for training (369 images) and validation (358). The target samples are only used for testing (361 images).

As features we use SIFT densely extracted from each image by using the VLFeat implementation [10]. We perform k-means clustering over a random subset of the descriptors from the source images and we build a visual vocabulary of 128 visual words. Every image is finally represented by a standard BOW feature vector.

We compare here the performance of three unsupervised domain adaptation techniques over the described experimental setup. The Geodesic Flow Kernel (GFK) method [5] defines a domain invariant feature representation by computing two domain specific subspaces for the source and

[1] http://homes.esat.kuleuven.be/~krematas/VisDA/CarEvolution.html

the target separately. They are then considered as points on a Grassmann manifold and all the data are projected into the intermediate subspaces along the shortest geodesic path connecting them. The Subspace Alignment (SA) approach introduced in [4] focuses instead on directly learning a mapping function which aligns the source subspace with the target one. Finally the Domain Adaptive Naive Bayes Nearest Neighbor (DA-NBNN) presented in [9] relies on NBNN and learns iteratively a class metric while inducing for each sample a large margin separation among the classes. Differently from the first two approaches, DA-NBNN runs over the SIFT descriptors avoiding the quantization step needed for the BOW features. For this last method we considered a random subselection of 300 SIFT vectors from each image.

All the results are shown in Figure 3. As baseline reference we present also the results obtained when learning on the source training set and testing both on the source validation set (SS), and on the target (ST) without adaptation. We used a linear SVM classifier for all the experiments, except DA-NBNN. The learning parameter C (100) was chosen by cross validation over the source data together with the BOW vocabulary dimension (128).

The divergence measure $H\Delta H$ [2] between the source and the target data demonstrates the existence of a domain shift: it is possible to discriminate the domains with an accuracy of about 70%, which explains also the recognition rate drop between SS and ST. From the GFK results we can state that this approach does not seem able to solve the domain mismatch, while both SA and DA-NBNN always improve over ST results, and outperform GFK to different extents. We point the interested readers to the corresponding reference papers for a more in-depth analysis of the methods.

4. Conclusions

In this paper we analyzed the problem of time domain shift for visual object classification. The design evolution of man-made objects makes any automatic classifier constantly outdated and rises the need of domain adaptation solutions. We introduced our car evolution dataset and we showed the performance of three state of the art adaptive methods over it. The proposed data testbed has been only marginally exploited. The results obtained over the source suggest that both the model and the manufacturer classification problems are challenging, in fact they correspond to difficult fine-grained tasks [1]. This indicates a direction for future research, towards algorithms able to integrate fine-grained and domain adaptation methods.

Acknowledgments We acknowledge the support from ERC grant Cognimund.

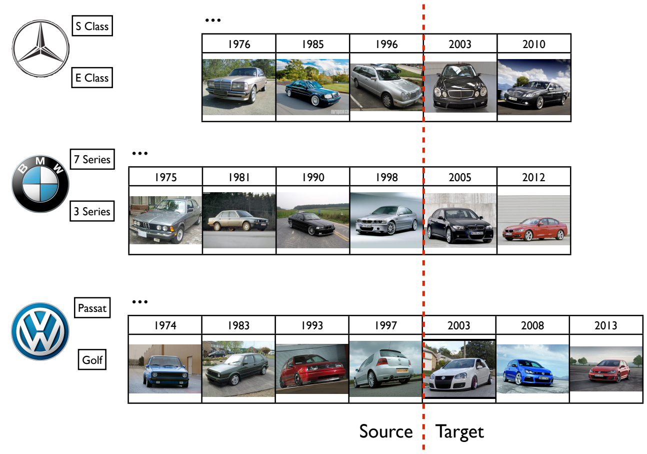

Figure 2. Car evolution hierarchy. Our dataset spans over 3 car manufacturers (Mercedes, BMW and VolksWagen (VW)), 6 models (E Class, S Class, 3 Series, 7 Series, Passat, Golf) and 35 lines in total. Note that here the years indicate the production date but not when the image was taken.

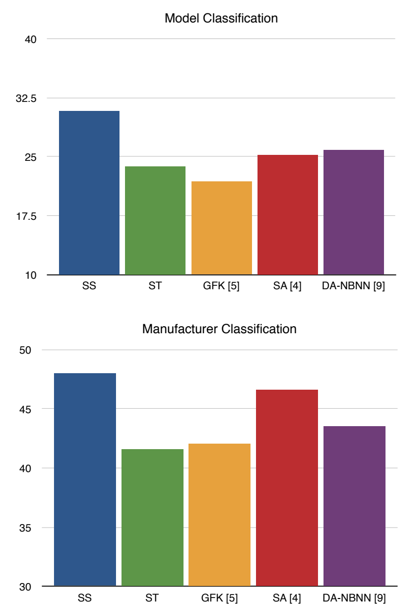

Figure 3. Recognition rate results (%) over the model and manufacturer car classification problem.

References

[1] The Fine Grained Challenge 2013. `https://sites.google.com/site/fgcomp2013/`.

[2] S. Ben-david, J. Blitzer, K. Crammer, and P. M. Sokolova. Analysis of representations for domain adaptation. In *NIPS*, 2007.

[3] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR*, 2009.

[4] B. Fernando, A. Habrard, M. Sebban, and T. Tuytelaars. Unsupervised visual domain adaptation using subspace align-

ment. In *ICCV*, 2013.

[5] B. Gong, Y. Shi, F. Sha, and K. Grauman. Geodesic flow kernel for unsupervised domain adaptation. In *CVPR*, 2012.

[6] Y. J. Lee, A. A. Efros, and M. Hebert. Style-aware mid-level representation for discovering visual connections in space and time. In *ICCV*, 2013.

[7] K. Saenko, B. Kulis, M. Fritz, and T. Darrell. Adapting visual category models to new domains. In *ECCV*, 2010.

[8] A. Shrivastava, T. Malisiewicz, A. Gupta, and A. A. Efros. Data-driven visual similarity for cross-domain image matching. *ACM SIGGRAPH ASIA*, 30(6), 2011.

[9] T. Tommasi and B. Caputo. Frustratingly Easy NBNN Domain Adaptation. In *ICCV*, 2013.

[10] A. Vedaldi and B. Fulkerson. VLFeat: An open and portable library of computer vision algorithms. `http://www.vlfeat.org/`, 2008.